APPENDIX 7B. TYPICAL METEOROLOGICAL YEAR (TMY2) ASSIGNMENTS

TABLE OF CONTENTS

LIST OF FIGURES

APPENDIX 7B. TYPICAL METEOROLOGICAL YEAR (TMY2) ASSIGNMENTS

7B.1 INTRODUCTION

7B.1.1 Purpose and Intent

This appendix describes the methodology by which *weighting factors* were developed to scale the influence of each TMY2 weather station on each state's results. The intent is to ensure that the cost-effectiveness calculations properly reflect the geographic distribution of buildings in which the HVAC systems under analysis will be used. State-level aggregate results need to be most heavily influenced by the climates having the most buildings.

7B.1.2 Issues

TMY2 stations, of which there are 239 in the United States, are distributed throughout the country so as to give good coverage of the climatic variation in the United States. Unfortunately for this study, the TMY2 stations are not distributed so as to match the distribution of buildings. It is therefore necessary to identify a reasonable mapping between the climate-based (TMY2) simulation results and the geographical distribution of buildings in each state. Although TMY2 data give good *climatic* coverage, they do not give sufficient *geographical* coverage.

Detailed data on the geographical/climatic distribution of buildings were not readily available for this task. However, data on the geographical distribution of persons (population data) can serve as a reasonable surrogate for buildings data, and are available from a number of sources. The problem facing this study was to match up population data, which are available in great geographical detail but not matched to climate indicators, with TMY2 data, which characterize climate quite well but are not connected to any population indicators.

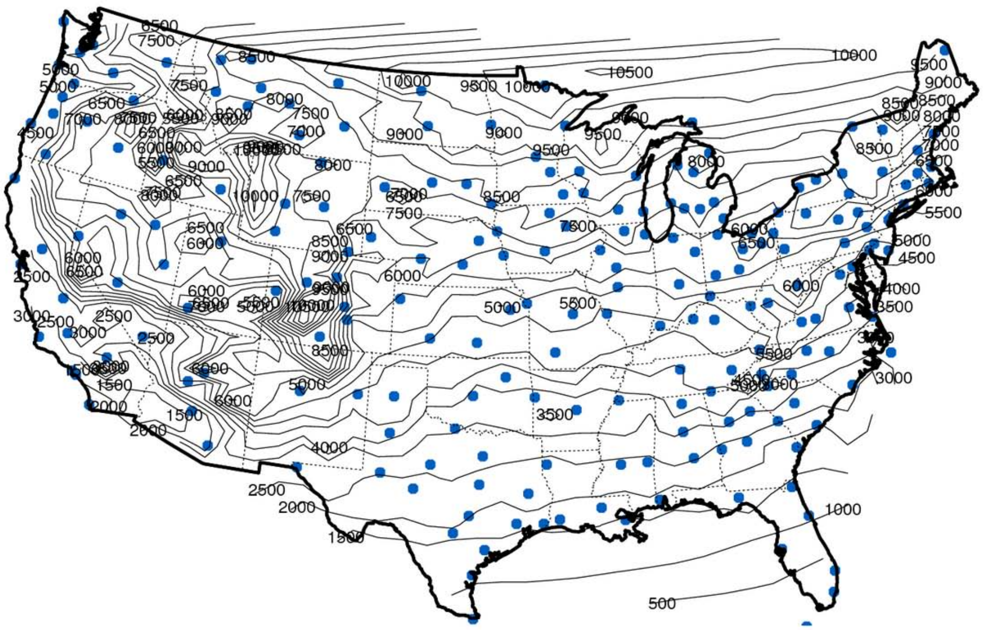
7B.2 APPROACH

Figure 7B.1 shows the locations of the 239 TMY2 stations superimposed on a contour map showing heating degree-days (HDD). It is clear that the geographic distribution of the TMY2 stations is too sparse to capture many of the climate variations that occur over relatively short distances, especially in mountainous regions. It is therefore not feasible to develop weighting factors based solely on the TMY2 stations within a state.

7B.2.1 Additional Data Sets

To develop weighting factors, we identified two data sets (in addition to the TMY2 data set) that provide a connection between detailed population distribution and TMY2 stations. These are:

- ***NOAA Climate Stations* –** The National Climatic Data Center's (NCDC)'s "CLIM81" database contains summary statistics from a large number of climate stations in the United States. We used the 1961–1990 period of record (POR), which corresponds to the POR used to define the TMY2 weather tapes, for which 4775 climate stations are represented (see http://lwf.ncdc.noaa.gov/oa/documentlibrary/clim81supp3/clim81.html). Each NOAA station is characterized by its location (latitude/longitude), annual heating and cooling degree-days, elevation, and various other metrics.



Points are TMY2 stations; contours are hdd65

Figure 7B.1 Locations of TMY2 Stations Superimposed on Heating Degree-Day Contours

- ***USGS Populated Places* –** The United States Geological Survey (USGS), in its Geographic Names Information System, maintains a detailed database of cities, towns, and other important features. (For the latest version of this data set, see http://geonames.usgs.gov/stategaz/00README.html. Our analysis used a version of this data set from the early 1990s, making it contemporaneous with the TMY2 and NOAA data sets.) Of interest to this task are the PPL features known as "populated places," which are generally cities and towns, but also include large housing subdivisions, trailer parks, and other places where people may live. For many of these populated places, the USGS has a population estimate. The populated places (PPL) data give excellent geographical coverage. The version of the data set used here has over 22,000 entries that include a population estimate out of more than 164,000 total.[1] Each PPL location is characterized by its location (latitude/longitude), elevation, population, and various other metrics.

The NOAA data are important because they contain climate summary information that can be mapped to the TMY2 stations, greatly increasing our geographical coverage. Figure 7B.2 shows the locations of the 4775 NOAA stations. The PPL database is used to make the final link between the climate information in the mapped NOAA/TMY2 stations and the geographical distribution of population. Figure 7B.3 shows the PPL locations.

1 Locations without USGS population estimates tend to be those with very little population. We have kept these smaller sites in our mapping analysis by assigning each a population estimate of one. This gives them negligible influence against the more populous locations but allows the mapping procedure to work even in very sparsely populated regions of some states.

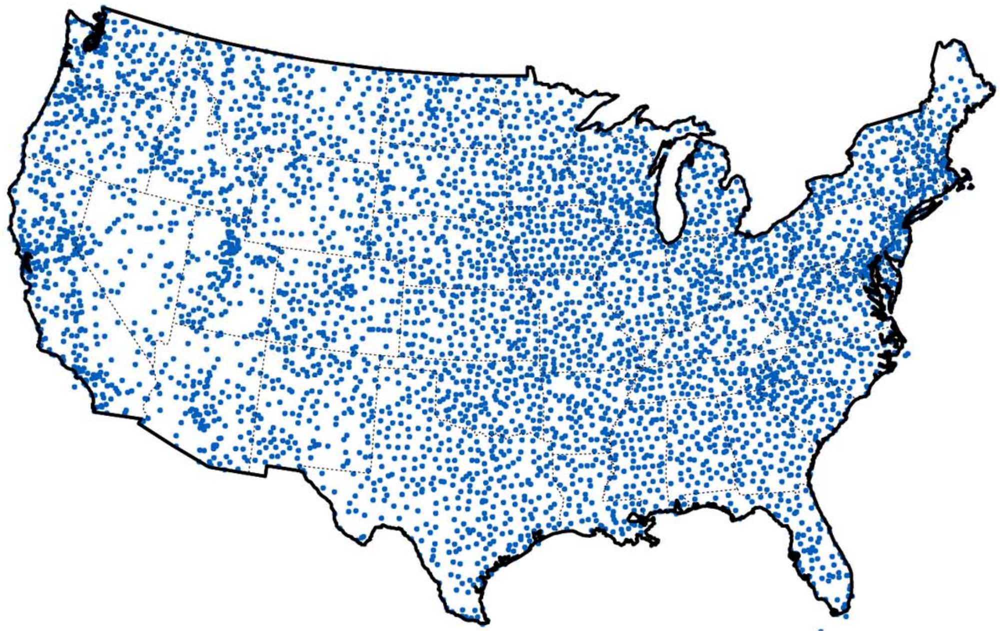


Figure 7B.2 Locations of 4775 NOAA Stations

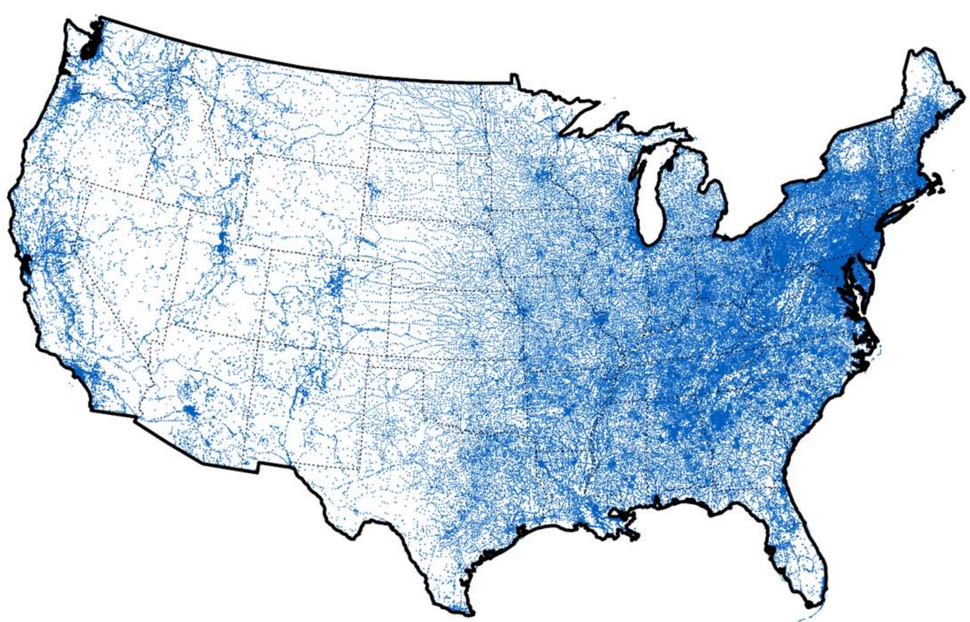


Figure 7B.3 Locations of Populated Places (PPL)

7B.2.2 Mapping Approach

Our approach to generating TMY2 weighting factors involves three major steps:

1. Map each PPL location in the United States to a best-representative NOAA station. This gives each location some summary climate metrics (chiefly, heating and cooling degree-days) that facilitate further mapping.

2. Map each NOAA station to a best-representative TMY2 station. This completes the link between the geographic population estimates (PPL data) and the TMY2 stations.

3. For each state, compute the fraction of the total PPL population that "points" (via its PPL→NOAA→TMY2 mappings) to each TMY2 station.

These are described in order below.

7B.2.2.1 Mapping *Populated Places* Locations to NOAA Stations

Mapping each of the 164,000+ PPL locations to a best-representative NOAA station is a mostly straightforward process. Because there is no climate information in the PPL database, the only metrics available to associate each PPL location with a NOAA station are location (latitude/longitude) and elevation (although elevation is not known for all PPL locations). The mapping algorithm is as follows.

1. For each PPL location, identify the nearest NOAA station. Distances between PPL/NOAA pairs are calculated using the latitudes and longitudes of the two locations and simple spherical geometry. If the elevation of the nearest NOAA station is within 300 feet of the PPL location or if the elevation of the PPL location is unknown, then the nearest NOAA station is the final mapping.

2. If the nearest NOAA station differs in elevation from the PPL location by more than 300 feet:

   a) Identify the 20 closest NOAA stations to the PPL location.

   b) Choose, from among the 20, the NOAA station that is nearest in elevation to the PPL location.

This algorithm is imperfect in many situations, but was designed by trial and error to give reasonable mappings in a large majority of cases. In locations with relatively flat terrain and fairly dense population, the algorithm almost always maps to the closest NOAA station. Figure 7B.4 shows the mappings in the state of Iowa as an example. Each plotted point on the graphic is one NOAA station; the "hairs" are drawn outward to the various PPL locations mapped to that station. In mountainous terrain or in locations with large distances between PPL locations (e.g., Alaska), the second part of the algorithm does a reasonable job of identifying a representative NOAA station, even if that station is some distance from the PPL location. Figure 7B.5 shows the mappings for Washington as an example. Note that many of the PPL locations (the ends of the hairs) are mapped to distant NOAA stations.

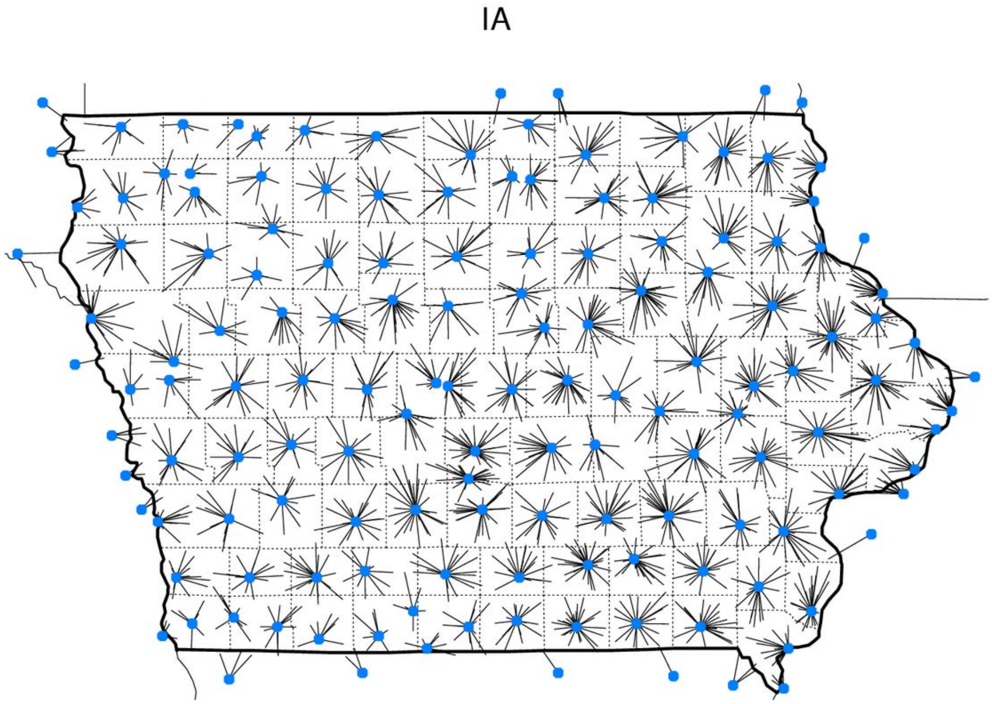


Figure 7B.4 PPL→NOAA Mappings for Iowa Showing a Predominance of Nearest-Location Mappings

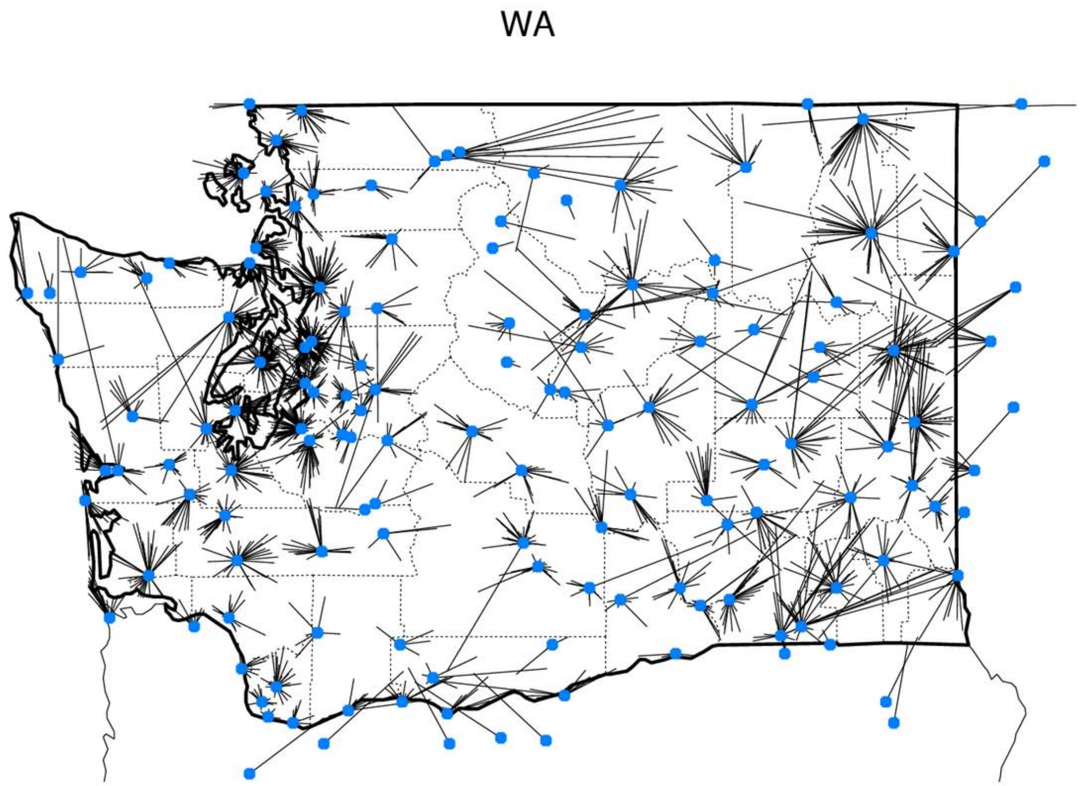


Figure 7B.5 PPL→NOAA Mappings for Washington Showing Some Distant Mappings in Mountainous Regions

7B.2.2.2 Mapping NOAA Stations to TMY2 Stations

Having mapped each PPL location to a best-representative NOAA station, the next step is to map each NOAA station to a best representative TMY2 station. Because the NOAA and TMY2 data sets provide less dense geographical coverage and because both data sets contain climate information, this mapping process relies less on spatial proximity and more on similarity of climate. Proximity remains important, though. In finding a best TMY2 mapping for a NOAA station, we hope for similarity in a number of climate variables: temperature (HDD and CDD), solar and wind characteristics, rainfall and humidity characteristics, snowfall, etc. The NOAA data set supports direct comparison only of heating and cooling degree-days, but we rely on spatial proximity to ensure similarity in the other variables.

The NOAA-to-TMY2 mapping algorithm is as follows.

1. For each NOAA station, identify the nearest TMY2 station. If the elevation of the nearest TMY2 station is within 300 feet of the NOAA station, then the nearest TMY2 station is the final mapping.
2. If the nearest TMY2 station differs in elevation from the NOAA station by more than 300 feet, then select instead the TMY2 station that has the minimum "combined distance" from the NOAA station. The combined distance is defined as the sum of the literal distance (miles) between the two stations and an "equivalent latitude miles" value that accounts for known differences in heating and cooling degree-days and elevation. This latter metric requires some explanation (see below).

The *equivalent latitude miles* metric was developed as a means to characterize temperature (HDD and CDD) and elevation in the same units as literal distance (i.e., miles). By casting HDD, CDD, and elevation effects into units of miles, we are able to simply sum the literal distance between two locations with the equivalent HDD/CDD/elevation distance to give the two values equal weight in assigning a best-representative TMY2 station to each NOAA station.

The equivalent latitude miles values are based on several related observations. First, differences in both heating and cooling degree-days correlate with differences in latitude. Second, those same degree-day differences also correlate with differences in elevation. Combining these observations, we discovered that differences in north-south distances (i.e., latitude miles) can be characterized in terms of differences in HDD, CDD, and elevation. That is, as one moves northward (increasing latitude) and upward (increasing elevation), HDD tends to increase and CDD tends to decrease. A regression analysis of these correlations allows us to cast HDD, CDD, and elevation differences into units of distance (miles). A linear regression on NOAA/TMY2 station pairs within 300 miles of each other gives the following result.

$$d_{equiv} = I + \alpha \times \Delta HDD + \beta \times CDD + \gamma \times \Delta Elev \qquad \textbf{Eq. 7B.1}$$

where

d_{equiv} = equivalent latitude distance (miles),
ΔHDD = difference in heating degree-days (base-65F)
ΔCDD = difference in cooling degree-days (base-65F)

$$\begin{aligned} \Delta Elev &= \text{difference in elevation (feet)} \\ I &= -6.8938 \\ \alpha &= 0.1061 \\ \beta &= -0.0149 \\ \gamma &= -0.0718 \end{aligned}$$

The best representative TMY2 station for each NOAA station was selected as the one with the minimum sum of actual distance and equivalent latitude distance. Figure 7B.6 shows the final NOAA→TMY2 mappings for the continental United States. Each plotted point is a TMY2 station, from which lines are drawn outward to the NOAA stations mapped to it.

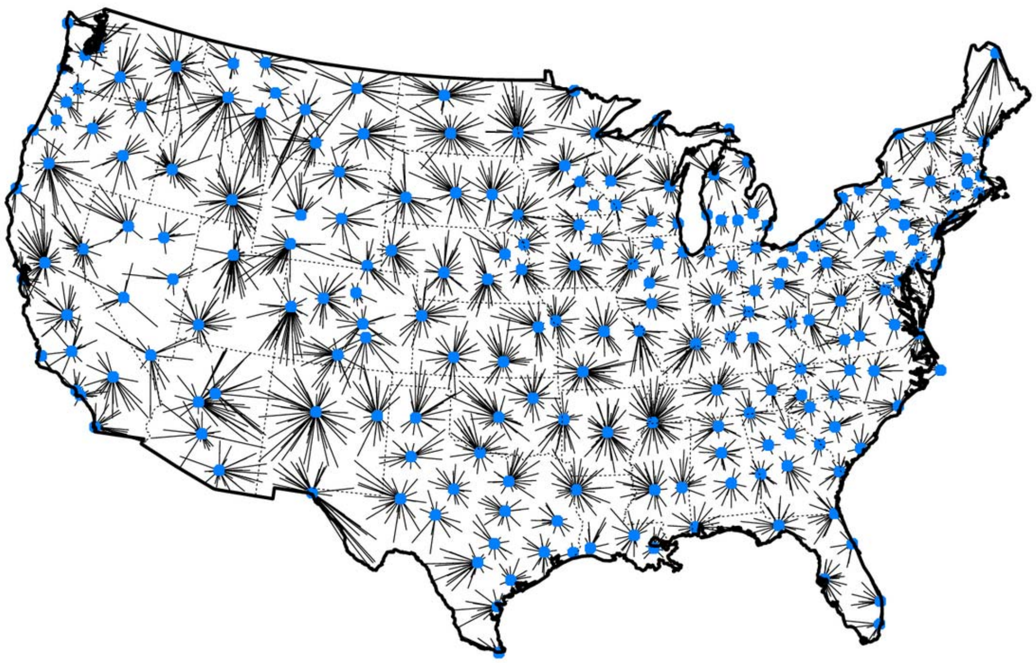


Figure 7B.6 NOAA→TMY2 Mappings

7B.2.3 Calculating State-by-State Weighting Factors

Weighting factors were developed for each U.S. state that express the fraction of that state's population that is represented by each TMY2 station. These weighting factors are based directly on the PPL→NOAA→TMY2 mappings described above. The weighting factor for a TMY2 station is defined as the summed population of all PPL locations in the state that point to that TMY2 station divided by the summed population of all PPL locations in the state. Thus, the sum of all the TMY2 weighting factors for a state is 1.0.